

多眼カメラを用いた超解像技術

著者	平尾 大樹
ページ	1-44
発行年	2016-03-24
学位授与年月日	2016-03-24
学位名	修士(工学)
学位授与機関	法政大学 (Hosei University)
URL	http://hdl.handle.net/10114/12523

2015年度(平成27年度)
修士論文

多眼カメラを用いた超解像技術

指導教員

彌富 仁

法政大学大学院 理工学研究科
応用情報工学専攻 修士課程

14R4124 ヒラオ ダイキ
 平尾 大樹

目次

	1
第1章 はじめに	2
第2章 目的・背景	3
第3章 方法	6
3.1 再構成型超解像技術の原理	6
3.2 位置合わせ処理	8
3.2.1 画像内の特徴点検出	9
3.2.2 透視投影変換による特徴点ベースの位置合わせ処理	12
3.2.3 Sum of Absolute Difference による画素値ベースの位置合わせ処理	16
3.2.4 Average Hash による類似度計算	18
3.3 再構成処理	19
3.3.1 画像劣化モデル	19
3.3.2 ML 法	19
3.3.3 最急降下法による最適化	20
3.4 複数のカメラを用いた多眼超解像システム	21
3.4.1 色補正処理	21
3.4.2 視差への対応	22
3.4.3 推定高解像度画像の雑音低減	22
3.5 多眼カメラを用いた超解像処理	23
3.6 超解像処理結果の数的評価	25
第4章 結果	26
4.1 近距離撮影画像に対する超解像処理	26
4.2 遠距離撮影画像を用いた超解像処理	29
4.3 暗所撮影画像を用いた超解像処理	32
4.4 それぞれの結果の数的評価	35
第5章 考察	37
第6章 まとめ	39
謝辞	40
参考文献	41
研究業績	43

第1章 はじめに

デジタルカメラ等を用いて高解像度画像を得るためには大きな撮像デバイスと優れた光学系が必要である。様々な技術革新と需要の増大によりシステムの低価格化は進んでいるものの、一般用途と高性能システムの間にはコストに大きな開きがある。また、画像観測の分野においても高解像度な画像の方がより詳細な画像情報を取得できるため、画像の高解像度化は常に求められている。このような問題に対して、観測した画像情報を基に高解像度画像を作成する超解像技術が提案されている。

超解像技術とは、低解像度の観測画像から高解像度の画像を作成する技術である。超解像技術は大きく2つの種類に分類することができる。1つは1枚の観測画像だけを利用した超解像技術であり、学習やデータベースに基づき高周波成分を復元する方法などがある。しかし、この方法で超解像を行う場合、生成された画像の高周波成分が真の高周波成分であるかはわからないという問題がある。もう1つは複数枚の観測画像を利用した超解像技術である。この方法は高い精度で超解像が実現できるという利点があるが、観測画像を大量に使用する必要がある。そのため、大量の画像を処理するのに計算時間がかかってしまうことや、観測時間にずれが生じてしまう動いている物体には処理が行えないというデメリットが存在する。

本研究では後者の超解像技術について取り組み、複数の安価なカメラをアレイ状に配置し、同期的に制御する「多眼超解像システム」を提案する。このようにして撮影を行う事により、観測時間にずれが生じていない観測画像を取得する事が可能である。しかし、この多眼超解像システムは各カメラの距離が少し離れているため、ピントのずれや視差によって観測画像ごとに違いが生じる可能性がある。そのため、本研究ではこのような多眼カメラの欠点を克服し、安価で動いている物体にも処理が可能な監視カメラの制作およびシステム開発を目的とする。

第2章 目的・背景

画像の解像度とは画像における画素の密度を示す数値であり、その密度が高ければ高いほど画像は鮮明に見える。超解像技術とは観測した低解像度画像を用いて高解像度な画像を生成する技術を指し、近年の間に超解像技術について数多くの研究が報告されている [1]。

これらは大きく2つに分けることができる。1つは学習型と呼ばれる超解像技術である。この手法は1枚の観測画像だけを用いて、予め学習したデータベースを参照することで高周波成分を復元する方法 [2] や、補間の考え方に基づく信号処理的なアプローチ [3] が提案されている。超解像とは、一般的に基となる観測画像に高周波成分が含まれていないと解像度の向上は見込めない。だが、この方法を使えば学習に基づき画像の高周波成分を付加できる可能性がある。しかし、その高周波成分が真の高周波成分かどうか分からないという欠点がある。

もう1つは再構成型と呼ばれる超解像技術である。この手法は複数枚の観測画像を用いて、処理後の画像の画素密度を高めていくことで解像度を向上させる手法である。この再構成型超解像の代表的なものとして、ML(Maximum-likelihood) 法、MAP(Maximum a posteriori) 法 [4][5]、IBP(Iterative backward projection) 法 [6]、周波数領域手法 [7]、POCS(Projection onto convex sets) 法 [8][9] などがあげられる。学習型とは違い、この再構成型は真の高周波成分を得ることができる。本研究では最終目的として多眼カメラによる監視カメラの制作を目指しているため、観測画像内の正しい情報を得ることが重要である。そこで、真の高周波成分を得ることができる再構成型の超解像処理を扱い、その中でも最尤推定の原理に基づくML法を使用する。ML法を用いた超解像処理の概要図を Fig.2.1 に示す。



Fig. 2.1: 超解像処理概要図

この再構成型超解像処理は複数枚の観測画像を使用するため、再構成処理の前処理として、観測の際の画像間の位置ずれを直す必要がある。画像間の位置ずれは平行移動のみとは限らず、回転や拡大縮小などの複雑な変形の場合がある。位置合わせ処理には、画像内の特徴点を抽出してその対応を利用する方法や、領域ベースマッチングなどのさまざまな手法 [10] がある。再構成処理による高画質化を実現するためには、サブピクセル精度の精密な位置合わせ処理が必要となり、超解像処理における課題の一つといえる。また、このような再構成型の超解像処理では観測画像全てに対して位置合わせ処理を行うため、処理時間が多くかかってしまう。本研究の位置合わせ処理には、主に SURF(Speeded Up Robust Features)[11] と透視投影変換を用いる。SURF による特徴点抽出と画像間の特徴点マッチングを利用して透視投影変換を行い、画像の位置ずれを推定する。

一般的に超解像処理は 1 台のカメラで撮影された観測画像を使用している。そのため、再構成型の超解像処理では撮影対象に動いている物体が存在していた場合、撮影画像ごとに時間のずれができてしまい、複雑な位置合わせ処理が必要となる。そこで、本研究では複数の低コスト撮像デバイスをアレイ状に配置し、同期的に制御するシステムを提案する。これにより、時間のずれが少ない複数の観測画像を取得できるため、動いている物体に対する高解像度化を望むことができる。また、低コストなカメラを使用しているため、高いコストをかけずに、高額なカメラを使用した場合と同等以上の画質を取得することも期待できる。カメラのセンササイズごとの価格の比較を Fig.2.2 に示す。本研究では Fig.2.2 の左側にあるよう

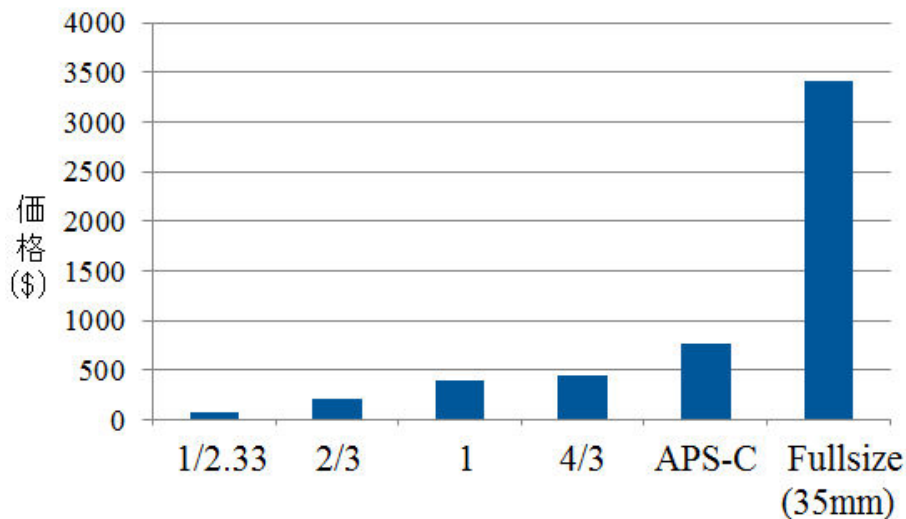


Fig. 2.2: カメラセンサと価格の比較

な安価なカメラを用い、Fullsize のような高額なカメラと同等以上の画質を取得することを目指す。

しかし、多眼カメラによる同時撮影を利用する場合、そのカメラの仕組みから様々な問題が生じてしまう。多眼カメラを用いて撮影した観測画像の例を Fig.2.3 と Fig.2.4 に示す。



Fig. 2.3: 多眼カメラによる観測画像 1



Fig. 2.4: 多眼カメラによる観測画像 2

多眼カメラは各カメラが少し離れているため、(1) カメラごとの色合いの違い、(2) カメラごとの視差、などの問題が生じてしまう。また、時間のずれが少ない画像を撮影できる枚数はカメラの台数までとなるため、(3) 撮影枚数がカメラの台数に依存する、といった欠点が挙げられる。そこで本研究では、これらの問題に対する対策も合わせて提案する。また、監視カメラでは様々な場面に対応する必要があるため、撮影対象までの距離が遠い場合、撮影対象までの距離が近い場合、撮影環境が暗い場合を想定し、実際に多眼カメラを用いて撮影をして処理を行う。

第3章 方法

3.1 再構成型超解像技術の原理

再構成型の超解像処理では、複数枚の観測画像を利用して高解像度画像を生成することができる。この概念を1次元画像を例にとりて考えてみる。仮に、1次元の画像2枚を観測し、その時のサンプル点が正確に半画素ずれている場合を想定する。このような場合、単純に2つの画像を結合する事により、Fig.3.1のように2倍の画素数の画像を得る事ができる。

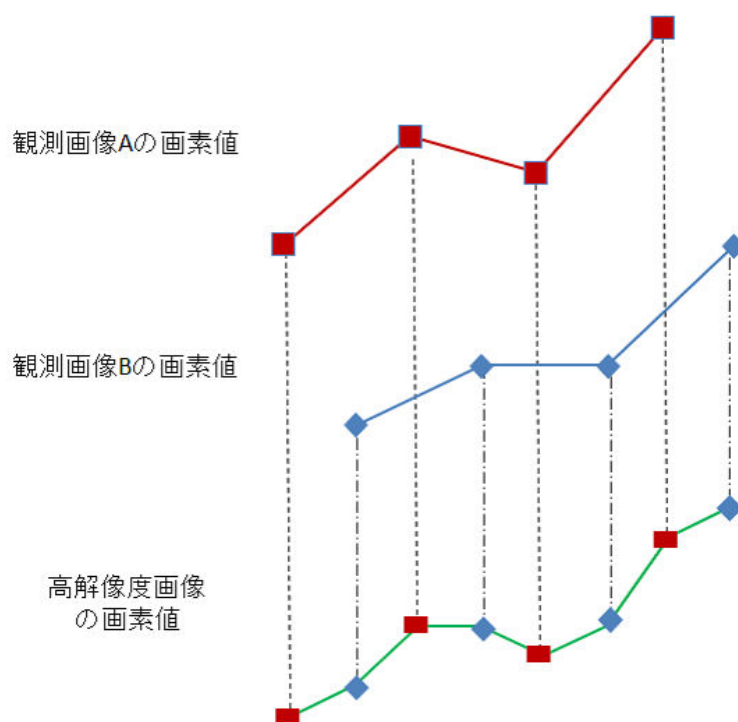


Fig. 3.1: 1次元の超解像処理の原理

2次元の画像の場合も同様に、サンプル点が半画素ずつずれている場合を想定すると、Fig.3.2のように4枚の観測画像を利用して 2×2 倍の画素数の画像を得る事ができる。

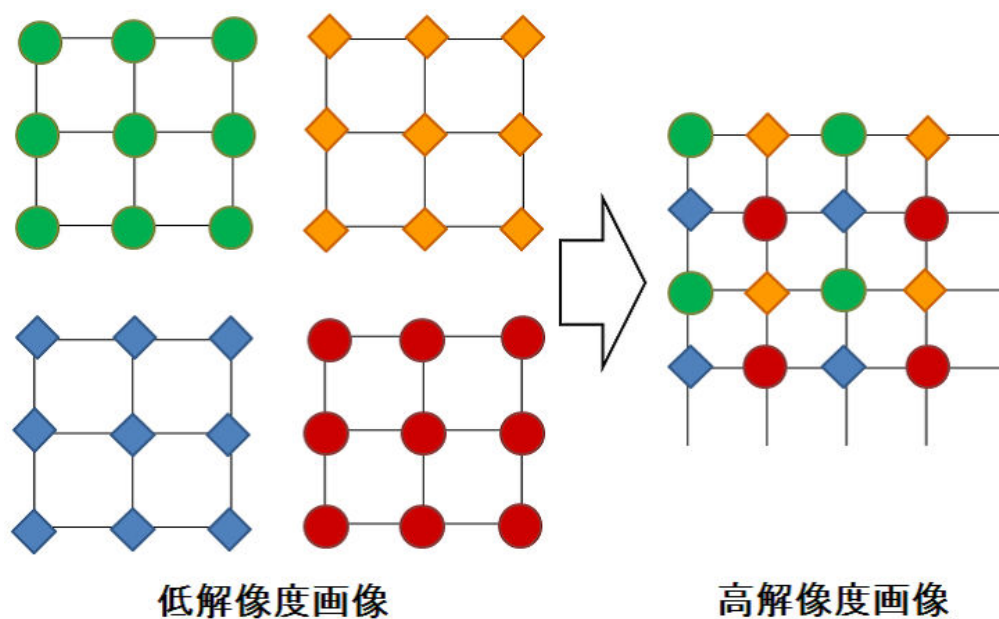


Fig. 3.2: 2次元の超解像処理の原理

このことから、超解像処理に必要な観測画像は、最低でも行いたい処理の倍率以上の枚数が必要であるとされている。また、このような考え方にに基づき、あらかじめ撮影デバイスを半画素ずつずらして配置し、高解像度画像を生成するシステム [12] も提案されている。

しかし、多くの場合で位置ずれを精密に制御する事はできず、複数枚の観測画像間の位置ずれは未知のことが多い。さらに、画像間の位置ずれは平行移動のみとは限らず、回転や拡大縮小などの複雑な変形の場合もある。そのため、再構成型の超解像処理ではサブピクセル精度の精密な位置合わせ処理が必要となる。

3.2 位置合わせ処理

上記のとおり、複数枚の観測画像を使用した再構成型超解像処理には、サブピクセル精度の正確な位置合わせ処理が必要になる。本研究では位置合わせ処理に画像の特徴点を利用しており、特徴点の検出には SURF や SIFT(Scale-Invariant Feature Transform)[13] を使用している。これにより、比較する 2 枚の画像のエッジ等から特徴点を検出する。その後、特徴点マッチングにより、2 枚の画像間で対応する点を取得する。そして、この対応点から 4 点を選び、透視投影変換を利用することで特徴点ベースの位置合わせ処理を行う。また、サブピクセル精度の位置合わせ処理を実現するため、2 枚の画像をピクセル精度で動かして比較することにより、画素値ベースの位置合わせ処理を行う。この 2 つの位置合わせ処理を組み合わせることにより、高精度な位置合わせ処理を実現する。また、これら 2 つの位置合わせ処理の後、Average Hash[14] という手法を用いて類似度を計算し、位置合わせ処理の精度を確認する。類似度が低いと判断された場合、透視投影変換に用いる SURF 特徴点を変更して、再度位置合わせ処理を行う。この時、繰り返し位置合わせ処理を行う回数は 4 回までとし、4 回目は SURF を用いていた特徴点検出を SIFT に切り替えて位置合わせ処理を行う。4 回とも位置合わせ処理を失敗してしまった場合、その画像は推定高解像度画像を更新するまでは使わないものとする。これらの位置合わせ処理の流れを Fig.3.3 に示す。

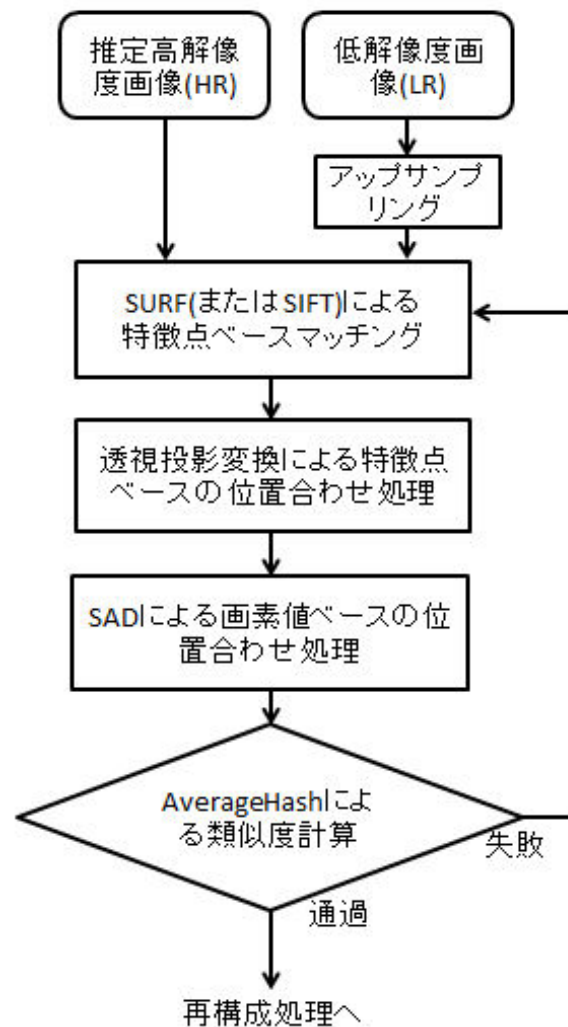


Fig. 3.3: 位置合わせ処理の流れ

以下にこれらについての概要を示す。

3.2.1 画像内の特徴点検出

画像内の特徴点の検出と特徴量の記述を行い、対応する点同士を結び付けることができるアルゴリズムに SURF というものがある。この SURF によって得られる特徴量は照明変化や回転、拡大縮小に不変であるという特徴がある。SURF は Hessian 行列という 2 次微分の集合を用いることで特徴点を検出している。Hessian 行列を (3.1) に示す。

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (3.1)$$

この L_{ab} はガウシアンの方方向の 2 次微分を画像に畳み込んだ応答値である。この行列が極大となる点を選ぶ事で、特徴点の検出を行う。しかし、このままではガウシアンを畳み込みに時間がかかってしまうため、(3.2) の矩形フィルタを用いて近似を行う。この様子を Fig.3.4 に示す。

$$\det(H_{approx}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (3.2)$$

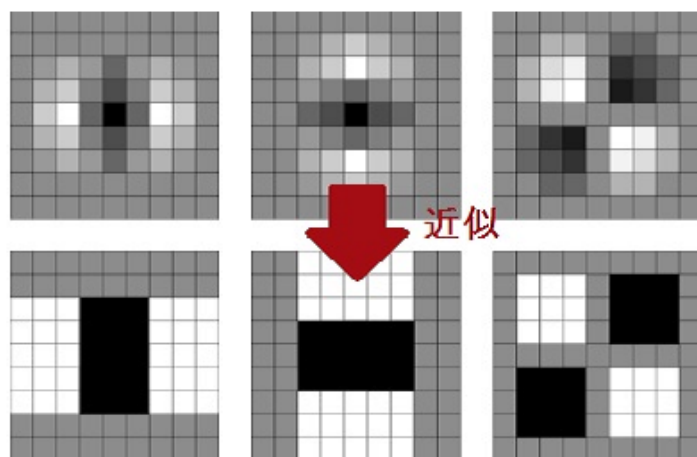


Fig. 3.4: 矩形フィルタによる近似

3.2 において 0.9 倍してあるのは、近似のために生じる誤差を修正するためである。これによって検出された特徴点は照明変化や回転、拡大縮小に頑強な特徴量であると言われている。

また、同じような特徴点検出のアルゴリズムに SIFT というものがある。SIFT も SURF と同様に、照明変化や回転、拡大縮小に不変な特徴量を取得する。しかし、この SIFT は SURF に比べて特徴点検出の精度が良いが処理に時間がかかるという欠点がある。再構成型超解像処理は撮影した大量の観測画像すべてに対して位置合わせ処理を行うため、精度を重視して SIFT のみを使うと処理に時間がかかってしまうという問題がある。下記に SURF および SIFT を用いた特徴点検出と特徴点マッチングの例を示す。以下の画像サイズ 640×360 の Fig.3.5 と Fig.3.6 対して SURF による特徴点検出と特徴点マッチングを行った結果を Fig.3.7 に示す。また、SIFT を用いた結果を Fig.3.8 に示す。各特徴点マッチング画像内の円は検出された特徴点を、線は 2 枚の画像の対応した特徴点同士を表している。



Fig. 3.5: 観測画像 A



Fig. 3.6: 観測画像 B

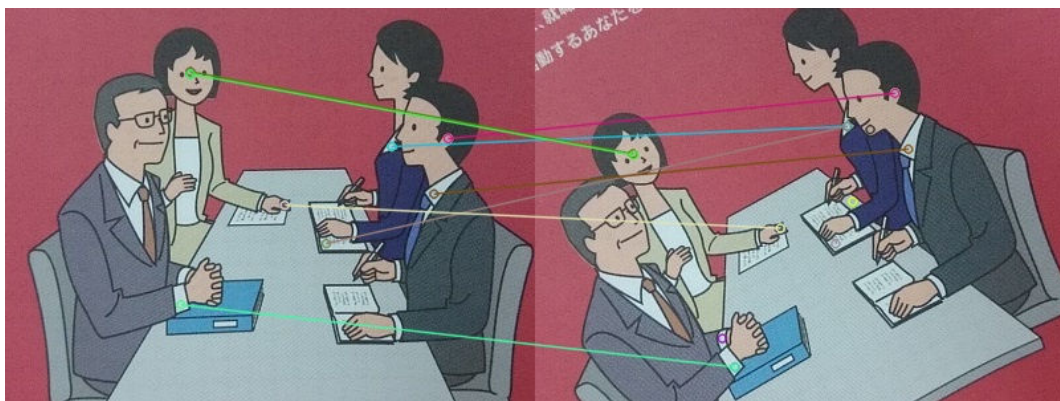


Fig. 3.7: SURF を用いた特徴点マッチング

Fig.3.7 と Fig.3.8 から、2 枚の画像間に対応している特徴点同士が結ばれている事がわかる。このとき、SURF によって検出した対応点は 8 点、処理時間は 0.39 秒である。また、SIFT



Fig. 3.8: SIFT を用いた特徴点マッチング

によって検出した対応点は 367 点、処理時間は 15.68 秒と、SURF と SIFT のどちらを使うかによってその精度や処理時間が大きく変わってくることがわかる。そのため、本研究では SURF と SIFT を組み合わせることにより、高精度かつ高速な位置合わせ処理を目指す。しかし、特徴点マッチングによって算出された対応している特徴点は、正確に一致している特徴点同士を算出できるとは限らない。そのため、イレギュラーな特徴点の組を削除する必要がある。そこで、本研究ではユーグリッド距離を用いて不要な特徴点を削除する。観測画像 A と観測画像 B の特徴点の座標をそれぞれ (x, y) 、 (x', y') として、特徴点の組の座標の変動量 σ を 3.3 で算出する。

$$\sigma = \sqrt{(x - x')^2 + (y - y')^2} \quad (3.3)$$

本研究では、変動量の大きい特徴点の上位 5 割を削除して使用する。

3.2.2 透視投影変換による特徴点ベースの位置合わせ処理

画像の平行移動や回転、拡大縮小などに対して、変換前後の4点の座標を指定して行う変換を透視投影変換という。透視投影変換を用いて2枚の画像に対して位置合わせ処理を行う場合、変換前と変換後でそれぞれ対応している4点の座標を指定する必要がある。そこで、2枚の画像に対しての特徴点マッチングを利用する。前述の特徴点マッチングを用い、画像間で対応した特徴点を透視投影変換における変換前の座標と変換後の座標と考えて処理を行い、位置合わせ処理を実行する。ここで、Fig.3.5とFig.3.6を例として、SIFTによって得られた特徴点を利用して透視投影変換を行う。はじめに、Fig.3.5を赤、Fig.3.6を青とした画像をFig.3.9に示す。



Fig. 3.9: 観測画像 A と B の比較

Fig.3.9 からわかるように Fig.3.5 と Fig.3.6 には大きな位置ずれがある。以下に、ランダムに選出した特徴点を用いて観測画像 B に対して透視投影変換を行った結果を示す。Fig.3.10、Fig.3.12、Fig.3.14 はランダムに選出された特徴点を結んだ線を表している。また、これらの特徴点を基にして透視投影変換を行った結果を Fig.3.11、Fig.3.13、Fig.3.15 に示す。



Fig. 3.10: ランダムに選出された特徴点 1



Fig. 3.11: 透視投影変換後の比較 1

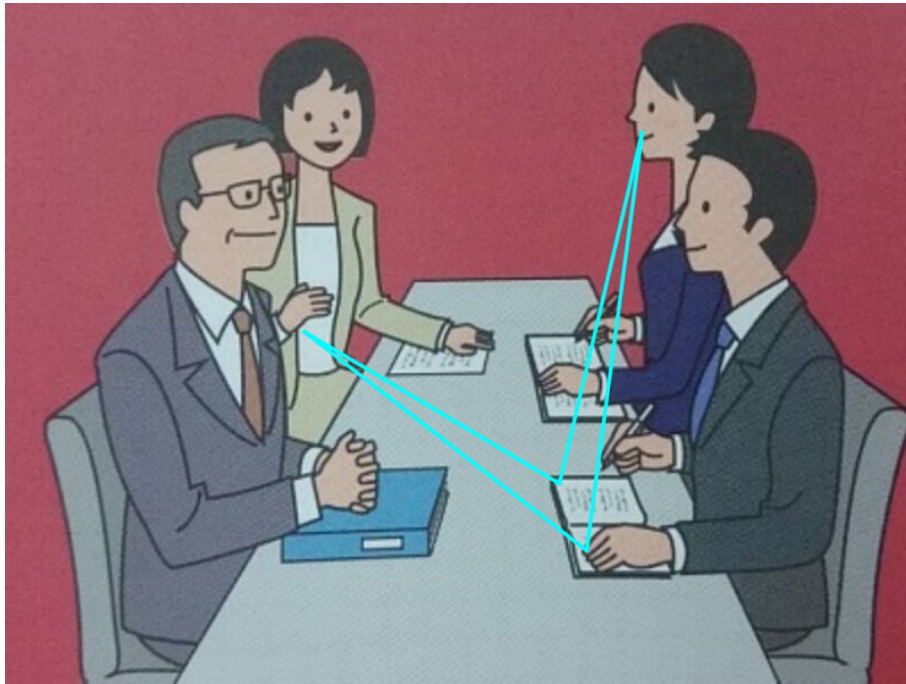


Fig. 3.12: ランダムに選出された特徴点 2



Fig. 3.13: 透視投影変換後の比較 2

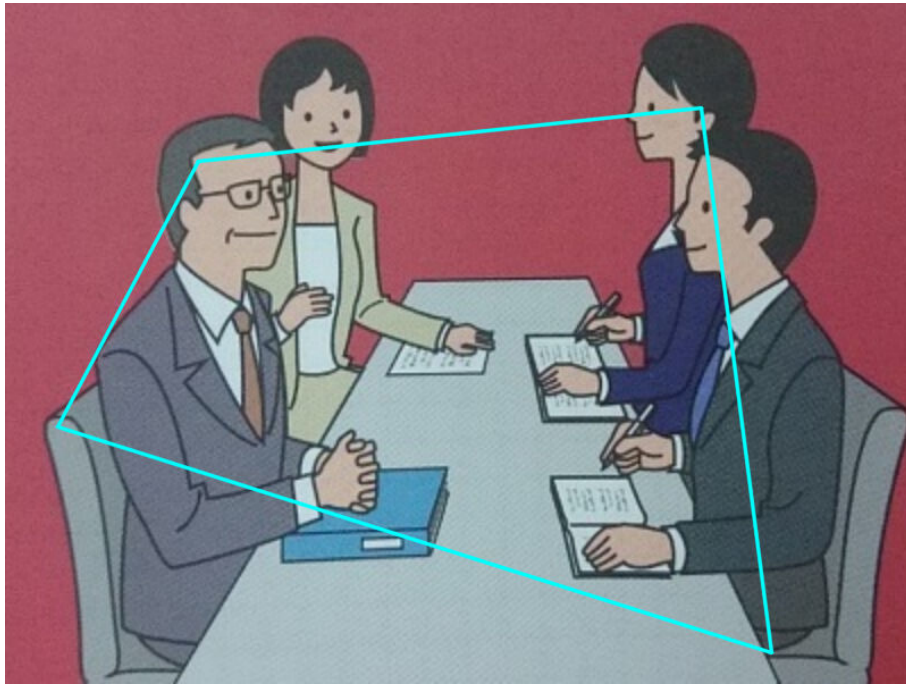


Fig. 3.14: ランダムに選出された特徴点 3



Fig. 3.15: 透視投影変換後の比較 3

以上の結果から、位置合わせ処理の精度が1番高いのはFig.3.14の結果ということがわかる。この事から、透視投影変換を行うときに選択する4点の座標は、4点を結んだ面積がより大きくなるように選択した方が正確な位置合わせ処理ができるのではないかと考えられる。そのため、本研究では透視投影変換を行うとき、残った信頼性の高い画像間の対応点の中で、結んだ面積がより大きく、またより画像全体をカバーできるように4点を選択する。具体的には画像の左上、左下、右上、右下の4隅から距離が近い特徴点をそれぞれ選択して処理を行うこととする。

3.2.3 Sum of Absolute Difference による画素値ベースの位置合わせ処理

透視投影変換による位置合わせ処理を行った後、さらなる位置合わせの精度向上のため、SAD(Sum of Absolute Difference) を利用した画素値ベースの位置合わせ処理を行う。SADとは輝度値の絶対値の差を算出する手法である。この計算方法を下記に示す。

$$R_{SAD} = \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} |I(i, j) - T(i, j)| \quad (3.4)$$

これを用い、2枚の画像の差を算出する。そして、位置合わせ処理を行う2枚の画像のうちの1枚を上下左右にピクセル精度で移動した画像を作成し、それぞれに対してSADの差を計算する。この時、算出された差の値が1番小さかったものがサブピクセル精度で正しい位置であるとする。5×5ピクセルの画像（Fig.3.16とFig.3.17）を用いて例を示す。ここで、各画像の太線で囲まれた部分が1ピクセル、太線内部の細い線が1本につき10の輝度値とする。

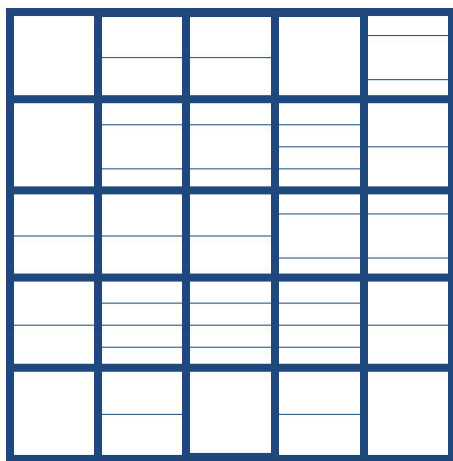


Fig. 3.16: 5×5ピクセルの画像例1

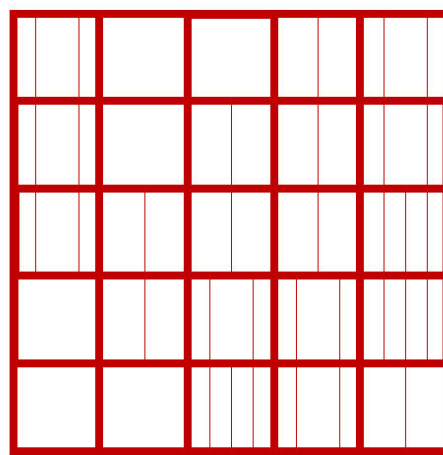


Fig. 3.17: 5×5ピクセルの画像例2

この例では上下左右に1ピクセルずつまで移動させることとする。Fig.3.16を基点とし、Fig.3.17を移動させた時の例をFig.3.18に示す。

画像例が5×5ピクセルに対し、1ピクセルずつまで移動させるので、計算する範囲は4×4ピクセルとする。また、2枚の画像が重なっている範囲が4×4ピクセル以上ある場合、重なっている部分の左上から4×4ピクセルを計算範囲とする。この計算範囲はFig.3.18内の黒の太枠で囲われている箇所である。

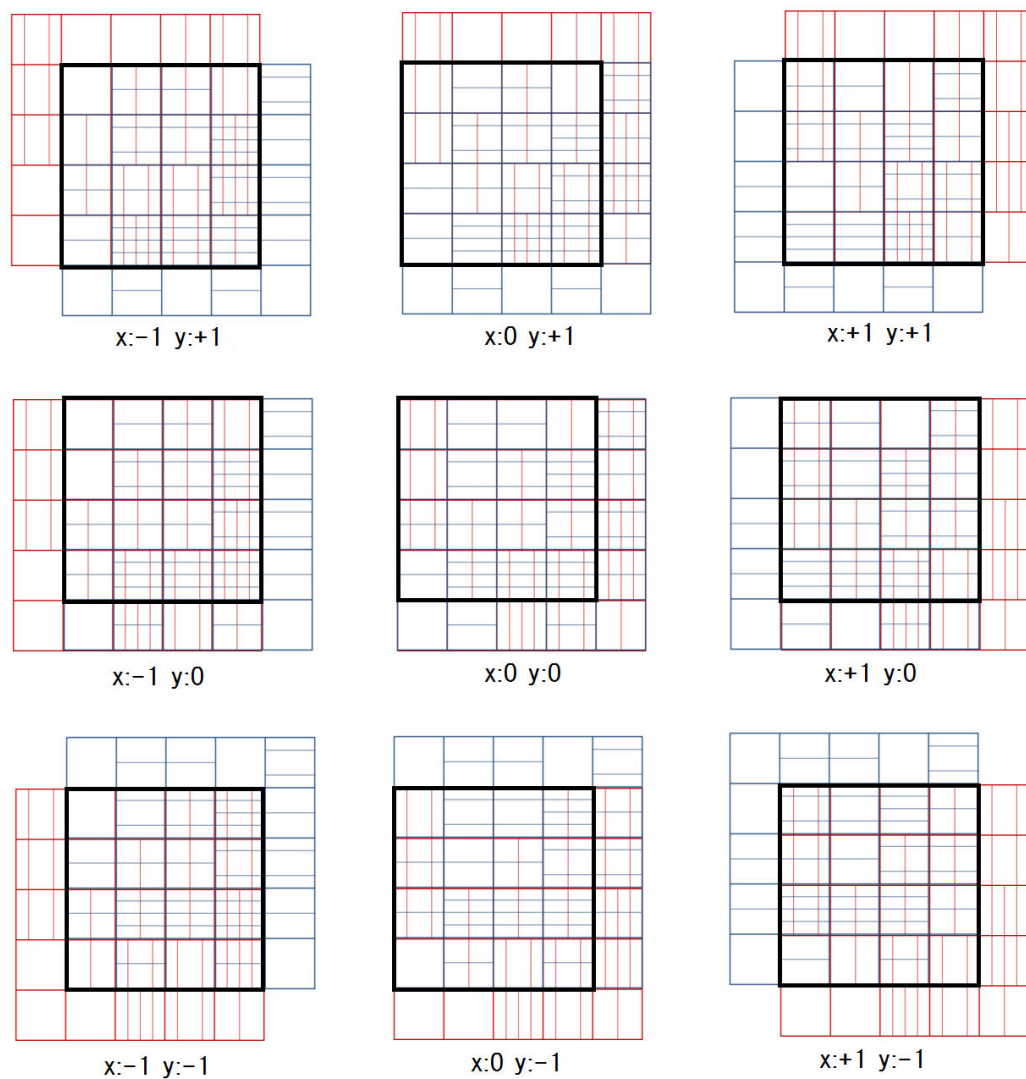


Fig. 3.18: SAD のピクセルシフトの例

輝度値の差が一番小さい例は、差が 90 となった X 軸方向に -1 ピクセルシフトさせた例であり、この状態が正しい位置であると考える。本研究の超解像処理の SAD では、上下左右に 3 ピクセルずつまで移動させて処理を行う。

3.2.4 Average Hash による類似度計算

上記の2つの位置合わせ処理を行った後、これらの処理の精度を確かめるために、Average Hash を用いて画像間の類似度を計算する。Average Hash は、グレースケール化した比較対象の領域画像を 8×8 ピクセルに縮小化し、各ピクセルの平均値を閾値として2値化することにより、 $8 \times 8 = 64$ ビットのビット列を生成する。そして、ビット列同士を比較することにより、画像間の類似度を高速に計算することができる手法である。このビット列への変換の流れを Fig.3.19 に示す。



Fig. 3.19: Average Hash の流れ

Fig.3.19 の場合では、生成されるビット列は1番下の行は「01111110」、下から2番目の行は「10111101」となる。本研究では精度を重視するために、予備実験の結果から、縮小化する正方形部位を 30×30 ピクセルに変更して類似度を求めた。

3.3 再構成処理

本研究の再構成処理では Farsin らの「Fast and Robust Multi-Frame Super-Resolution」[15] を参考に作成している。手法としてはカメラで撮影する際の観測画像が劣化する原因を考え、ML 法を用いて高解像画像を生成する手法である。以下に再構成処理に使用するこれらについての概要を示す。

3.3.1 画像劣化モデル

はじめに、画像の劣化モデルについて考える。もともと高画質だった撮影対象が観測の際に劣化してしまうのは、その撮影過程でさまざまな影響を受けてしまうためであると考えられる。この劣化の原因として、主に大気の影響や手振れ、カメラの汚れ、保存の際のダウンサンプリングなどが挙げられる。この画像劣化モデルは (3.5) のように表現できる。

$$Y_k = D_k H_k F_k X + V_k \quad k = 1, \dots, N \quad (3.5)$$

Y は観測した低解像度画像、 D はダウンサンプリング、 H はぼけ関数、 F は動きによる影響、 X は推定高解像度画像、 V は雑音、 N は入力画像の枚数を表している。

3.3.2 ML 法

不明な高解像度画像を、観測した低解像度画像と上記に示した画像劣化モデルに基づき ML 法を用いて推定する。ML 法とは最尤推定の原理に基づく再構成型の超解像処理の代表的な手法である。推定高解像度画像を母数、そこから得られると考えられる観測低解像度画像を標本とし、観測低解像度画像から未知数である推定高解像度画像が得られる確率が最大になるような推定高解像度画像を探索する。これを最小化問題として書き直すと (3.6) のようになる。

$$X = \arg \min_X \left[\sum_{k=1}^N \rho(Y_k, D_k H_k F_k X) \right] \quad (3.6)$$

具体的な数式として、(3.7) に示すように推定した高解像度画像に前述の劣化の影響を加えたものと観測低解像度画像との誤差を算出して、その誤差が最小になるような高解像度画像を推定する。

$$X = \arg \min_X \left[\sum_{k=1}^N \|D_k H_k F_k X - Y_k\|_p^p \right] \quad (3.7)$$

また、式中の p は最小誤差問題の L ノルムについての記述であり、本研究では $p = 1$ 、つまり L_1 ノルムを扱う。

$$X = \arg \min_X \left[\sum_{k=1}^N \|D_k H_k F_k X - Y_k\|_1 \right] \quad (3.8)$$

3.3.3 最急降下法による最適化

(3.8) の最小化問題を解決するため、最急降下法に基づいて式を書き直すと下記のようになる。

$$X_{n+1} = X_n - \left\{ \sum_{k=1}^N F_k^T H_k^T D_k^T \text{sign}(D_k H_k F_k X_n - Y_k) \right\}. \quad (3.9)$$

ここで、 F は推定高解像度画像を低解像度画像の位置に合わせる位置合わせ処理、 F^T は位置合わせした画像を元の位置に戻す処理、 H^T はぼかし処理を戻す反ぼかし処理、 D^T はダウンサンプリングを行った分のアップサンプリング、 sign は符号関数であり、正の場合は1を、負の場合は-1を返す関数を表している。

3.4 複数のカメラを用いた多眼超解像システム

再構成型の超解像処理は、大量の観測画像を使用する。しかし、1台のカメラを用いた連続撮影によって得られる画像は、どうしても撮影時間にずれが生じてしまう。この時、撮影される物体が静止しているのなら問題はないが、撮影対象に動いている物体が存在していた場合、観測画像ごとに位置がずれてしまうため、その事を考慮した複雑な位置合わせ処理が必要となる。そこで、本研究では複数のカメラをアレイ状に配置し、同期的に制御する事で観測画像を取得する。撮影に使用する装置のプロトタイプを Fig.3.20 に示す。



Fig. 3.20: 多眼カメラ

このプロトタイプは12個のカメラ (AXIS M1011-W) から構成されていて、各カメラの解像度は 640×480 ピクセルであり、各カメラの視野角は水平方向が 47° 、垂直方向が 35° である。また、プロトタイプ全体の大きさは $18\text{cm} \times 24\text{cm}$ である。これによって、撮影時間のずれがないほぼ同視野の観測画像を複数取得できるため、動いている物体に対しても超解像処理を行うことが可能となる。しかし、多眼カメラはアレイ状に配置されているため、観測画像ごとに視差の違いやピントのずれによる色合いの違いが起こる。そのため、本研究では多眼カメラによって撮影された画像を用いて超解像処理を行うために、2つの前処理および画質改善策の計3つのアイデアを提案する。

3.4.1 色補正処理

本研究では多眼カメラによって観測画像を取得している。しかし、多眼カメラは各カメラごとに距離が少しずつ離れているため、カメラごとに少しピントがずれてしまう可能性がある。カメラはピントの位置によって観測画像の色合いが変わってくるので、多眼カメラによって得られた観測画像は画像ごとに色合いが違う可能性がある。色合いが違う観測画像をそのまま超解像処理に用いてしまうと、高画質化に悪影響を及ぼしてしまう可能性がある。

そこで、任意に選択した1枚の観測画像（推定高解像度画像として扱う）を基準として、全ての観測画像に対してヒストグラムマッチングによる色補正処理を行う。これにより、観測画像ごとの色合いを合わせてから超解像処理を行うことで、より高い解像度を望むことができると考えられる。

3.4.2 視差への対応

多眼カメラではカメラ間で視差が発生し、これは撮影対象までの距離が近い場合により顕著に表れる。視差がある画像同士は、物体の奥行きなどの影響をうけ、正確な位置合わせ処理は困難となるため、精密な位置合わせ処理が要求される超解像処理では視差の処理が重要になる。そのため、本提案システムでは解像度を高くしたい部分 (ROI) だけを半自動的に切り取って処理を行う方法を提案する。まず初めに、任意に選択した i 番目の観測画像 ($1, 2, \dots, i, \dots, N$) から ROI_i を手動で切り取る。カメラによって撮影された観測画像を Fig.3.21 に、切り取った観測画像を Fig.3.22 に示す。



Fig. 3.21: 撮影した観測画像



Fig. 3.22: 手動で切り取った観測画像

この切り取った観測画像を基に、SURF による特徴点マッチングを使用して画像間に対応する領域 $ROI_j (1, 2, \dots, j, \dots, N, j \neq i)$ を自動的に抽出する。切り取った観測画像と他の観測画像に対し、特徴点マッチングを行い、信頼度の高い対応点のみを検出する。この時の様子を Fig.3.23 に示す。

この信頼度の高い対応点を基に、手動で切り取った画像と同じ個所を自動で切り取っていく。本研究ではこうして得られた画像を超解像処理における入力画像として扱うことで、視差の大きくなってしまった観測画像に対しての処理を行う。

3.4.3 推定高解像度画像の雑音低減

通常、低コストカメラは小さいセンサー ($1/4$ 、 $1/2.33$ など) を使用している。これらのセンサーを使用する場合、照明が不十分な環境で撮影を行うと、観測画像には雑音が付加されてしまうことが多い。また、多眼カメラ超解像システムは、観測画像の枚数がカメラの台数に依存するという欠点がある。使用する観測画像が少ない状態で超解像処理を行うと、画像内のエッジなどと同時に雑音も強調してしまう。そのため、以下の2つの処理を超解像プロセスに組み込むことを提案する。



Fig. 3.23: ROI 自動抽出のための特徴点マッチング

本研究の超解像処理では、観測画像の1枚をアップサンプリングしたものを初期の推定高解像度画像として使用している。この時、まだ再構成処理を行っていない推定高解像度画像は、他の観測画像と同様に雑音が付加されている。そのため、従来の再構成型の超解像プロセスの初回にメディアンフィルタを用いて推定高解像度画像の雑音除去を行うことを提案する。また、超解像処理に使用する観測画像が少ない場合、ML法による推定高解像度画像の更新処理と同時に雑音が付加されてしまう可能性がある。そのため、本研究では推定高解像度画像と観測画像との誤差に対してバイラテラルフィルタを使用することによって、エッジを保持しつつ雑音を除去してから更新処理を行う。これらの方法を用いることによって、利用する低解像度画像の枚数が少ない場合でも、画像内の輪郭をぼやかすことをせずに雑音を除去することが期待できる。

3.5 多眼カメラを用いた超解像処理

以上のことを踏まえ、超解像プログラムを作成する。今回作成した超解像処理プログラムのモデル図を Fig.3.24 に示す。

Fig.3.24 の緑で囲まれた部分は観測画像を表し、そのうちの1枚をアップサンプリングした画像を最初の推定高解像度画像として扱い、それ以外の画像を入力低解像度画像として使用する。また、位置合わせ処理には Fig.3.3 を使用している。本研究では H_k と H_k^T の PSF にはフーリエ変換を用いたローパスフィルタとハイパスフィルタを利用し、切り取る周波数とその減衰域を変化させる事で調整を行う。

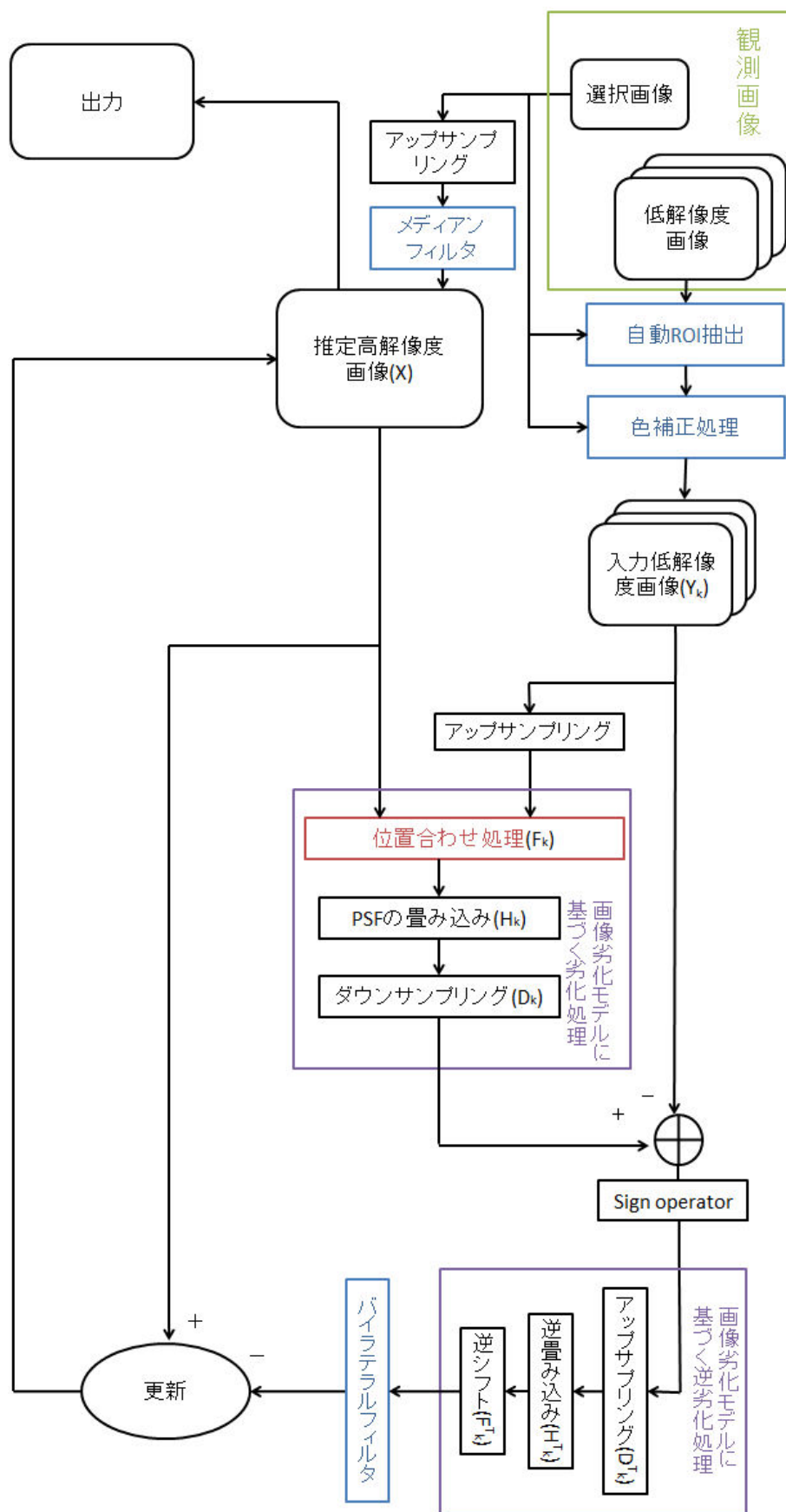


Fig. 3.24: 作成した超解像処理のモデル図

3.6 超解像処理結果の数的評価

超解像処理では取得した画像の数倍の画質の画像を生成することが目的であるため、ゴールドスタンダードとなるデータは存在しない。そこで、本研究では多眼カメラによる連続撮影を利用し、それによって得られた観測画像 40 枚を用いた従来手法の超解像処理結果をゴールドスタンダードとして扱う。Fig.3.22 に対するゴールドスタンダードの結果を Fig.3.25 に示す。この画像を真の結果し、S/N 比によって従来手法および提案手法の数的評価を行う。

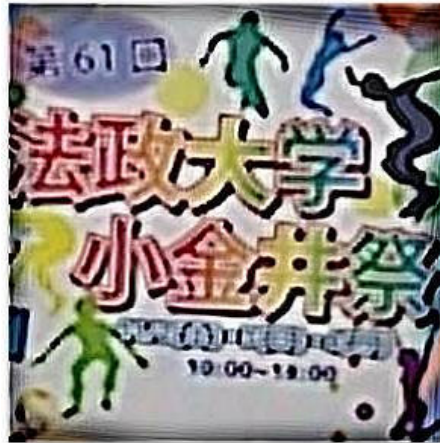


Fig. 3.25: 40 枚の観測画像を利用した従来 of 超解像処理結果

第4章 結果

多眼カメラによって撮影した観測画像を用いて、作成した超解像処理を確認した。この時、処理を行った倍率は 3×3 倍である。

4.1 近距離撮影画像に対する超解像処理

まず初めに、近距離を撮影した Fig.3.22 に対して超解像処理を行う。bi-linear 法を用いて拡大した結果を Fig.4.1 に、観測画像 12 枚を用いた提案超解像手法の結果を Fig.4.2 に示す。また、これらの拡大画像を Fig.4.3、Fig.4.4 に示す。



Fig. 4.1: 近距離撮影画像を用いた bi-linear 法 Fig. 4.2: 近距離撮影画像を用いた提案超解像手法 (12 枚)



Fig. 4.3: 近距離撮影画像を用いた拡大bi-linear 法



Fig. 4.4: 近距離撮影画像を用いた提案超解像手法（12枚）

これらの結果から、提案する色補正処理や自動ROI抽出が上手く働き、視差が大きな近距離を撮影した観測画像に対しても超解像処理が行えていることがわかる。また、観測画像が全てのカメラで上手く撮影できなかった状態、多眼カメラの小型化のために使用するカメラの台数を減らした状態等を想定した、観測画像の枚数が少ない場合（4枚）の結果を Fig.4.5 から Fig.4.8 に示す。



Fig. 4.5: 近距離撮影画像を用いた従来手法（4枚）



Fig. 4.6: 近距離撮影画像を用いた提案手法（4枚）



Fig. 4.7: 近距離撮影画像を用いた拡大従来手法（4枚）

Fig. 4.8: 近距離撮影画像を用いた拡大提案手法（4枚）

近距離を撮影した観測画像を用いた場合では、従来の手法では観測画像が少ない状態で超解像処理を行うと雑音も強調してしまうが、提案手法では雑音が低減されていることがわかる。

また、他の撮影環境も想定し、遠距離、暗所での観測画像を用いて超解像処理を行う。それぞれの結果は観測画像の枚数がぎりぎり足りている場合（9枚）と、観測画像の枚数が少ない場合（4枚）の2種類を示す。

4.2 遠距離撮影画像を用いた超解像処理

遠距離を撮影した観測画像を用いて処理を行い、その結果を Fig.4.9 から Fig.4.13 に示す。それぞれ bi-linear 法を用いた結果 (Fig.4.9)、9 枚の観測画像を用いた従来超解像手法の結果 (Fig.4.10)、および提案手法の結果 (Fig.4.11)、4 枚の観測画像を用いた従来手法の結果 (Fig.4.12)、および提案手法の結果 (Fig.4.13) である。また、それぞれの一部拡大画像を Fig.4.14 から Fig.4.18 に示す。



Fig. 4.9: 遠距離撮影画像を用いた bi-linear 法



Fig. 4.10: 遠距離撮影画像を用いた従来超解像手法 (9 枚)



Fig. 4.11: 遠距離撮影画像を用いた提案超解像手法 (9 枚)



Fig. 4.12: 遠距離撮影画像を用いた従来超解像手法 (4 枚)



Fig. 4.13: 遠距離撮影画像を用いた提案超解像手法 (4 枚)



Fig. 4.14: 遠距離撮影画像を用いた拡大 bi-linear 法



Fig. 4.15: 遠距離撮影画像を用いた拡大従来手法 (9 枚)



Fig. 4.16: 遠距離撮影画像を用いた拡大提案手法 (9 枚)



Fig. 4.17: 遠距離撮影画像を用いた拡大従来手法（4枚）



Fig. 4.18: 遠距離撮影画像を用いた拡大提案手法（4枚）

遠距離を撮影した観測画像を用いた場合では、従来手法と提案手法に視覚的に優位な差はあまり見れなかった。

4.3 暗所撮影画像を用いた超解像処理

最後に、暗所を撮影した観測画像を用いて処理を行う。この時、撮影した暗所の観測画像に対し、ヒストグラムイコライザによって画像を見やすくしてから超解像処理を行った。この時の超解像処理結果を Fig.4.19 から Fig.4.23 に示す。それぞれ bi-liner 法を用いた結果 (Fig.4.19)、9 枚の観測画像を用いた従来手法の結果 (Fig.4.22)、および提案手法の結果 (Fig.4.23)、4 枚の観測画像を用いた従来手法の結果 (Fig.4.20)、および提案手法の結果 (Fig.4.21) である。また、それぞれの一部拡大画像を Fig.4.24 から Fig.4.28 に示す。



Fig. 4.19: 暗所撮影画像を用いた bi-linear 法



Fig. 4.20: 暗所撮影画像を用いた従来超解像手法 (9 枚)



Fig. 4.21: 暗所撮影画像を用いた提案超解像手法 (9 枚)



Fig. 4.22: 暗所撮影画像を用いた従来超解像手法 (4 枚)



Fig. 4.23: 暗所撮影画像を用いた提案超解像手法 (4 枚)



Fig. 4.24: 暗所撮影画像を用いた拡大 bi-linear 法



Fig. 4.25: 暗所撮影画像を用いた拡大従来手法 (9 枚)



Fig. 4.26: 暗所撮影画像を用いた拡大提案手法 (9 枚)



Fig. 4.27: 暗所撮影画像を用いた拡大従来手法 (4 枚) Fig. 4.28: 暗所撮影画像を用いた拡大提案手法 (4 枚)

暗所を撮影した観測画像を用いた場合では、従来手法を用いた結果は雑音も強調されてしまっている。しかし、提案手法ではその雑音が低減されていることがわかる。

4.4 それぞれの結果の数的評価

従来手法と提案手法の数的評価のために、40枚の観測画像を利用した超解像処理の結果を真と仮定した時、使用した観測画像の枚数と超解像手法によるSN比の変化を Fig.4.30 から Fig.4.29 に示す。

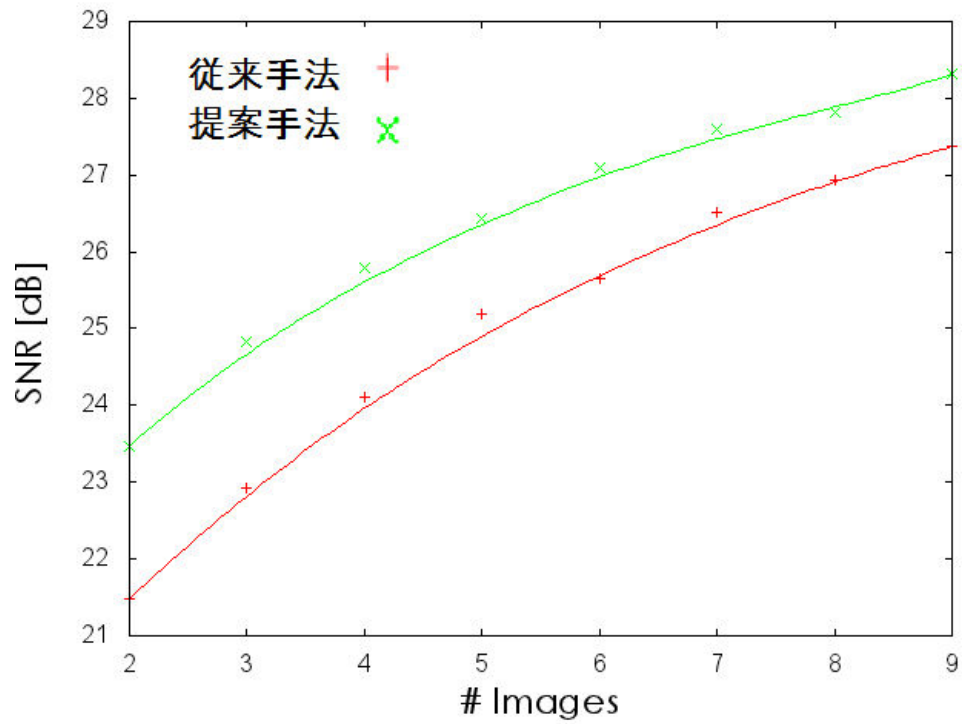


Fig. 4.29: 近距離撮影画像を用いた超解像処理結果の S/N 比

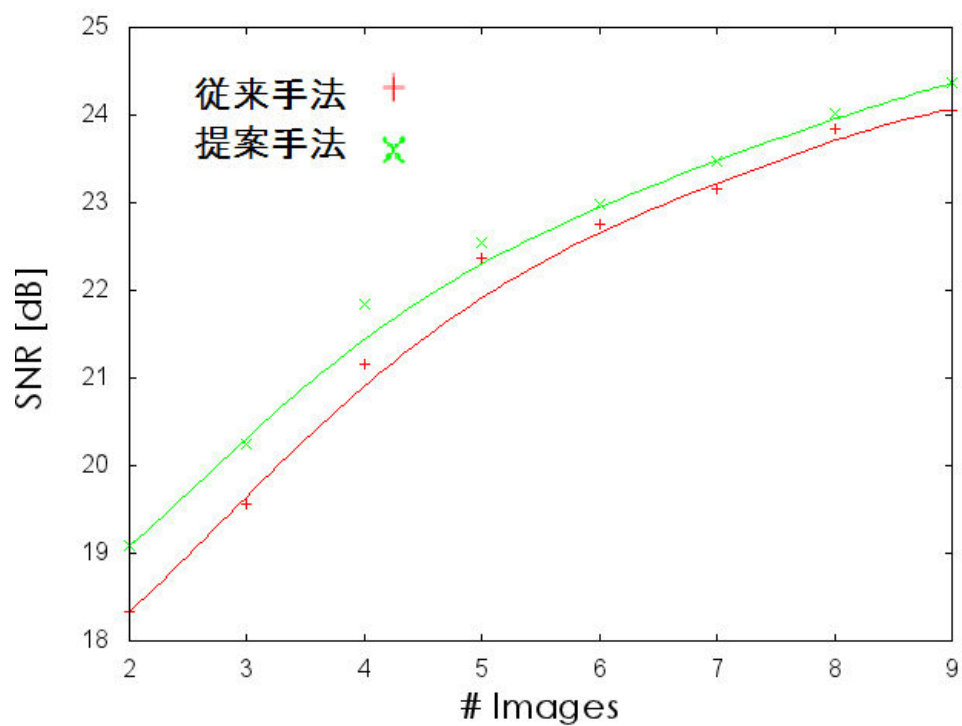


Fig. 4.30: 遠距離撮影画像を用いた超解像処理結果の S/N 比

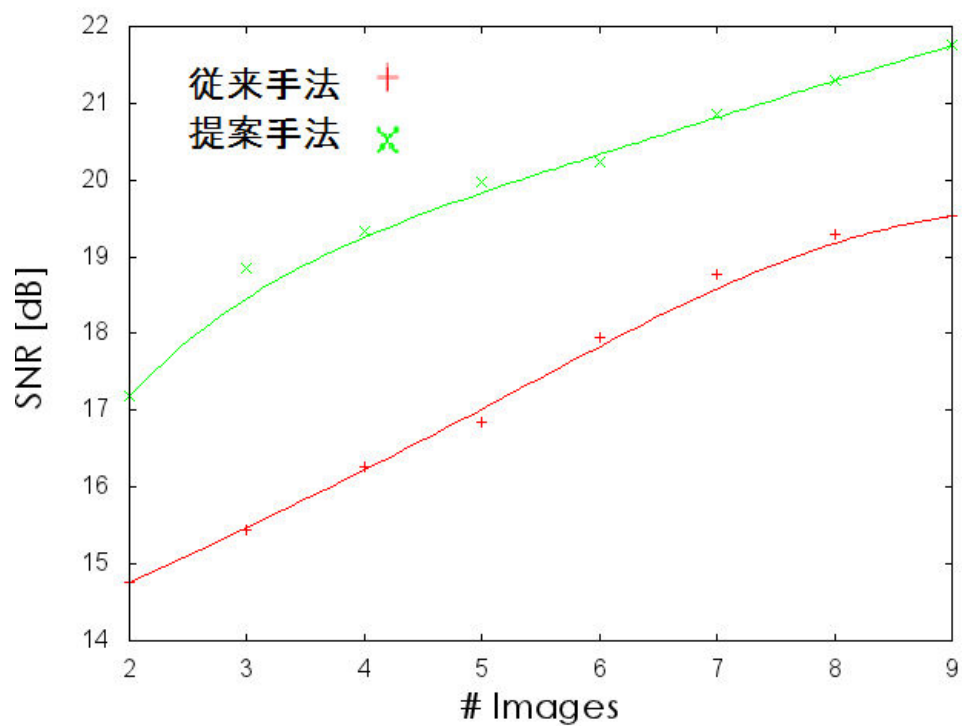


Fig. 4.31: 暗所撮影画像を用いた超解像処理結果の S/N 比

以上の結果から、撮影環境によって差はあるものの、従来手法に比べて提案手法の結果の方が雑音が少ないことがわかる。

第5章 考察

まず初めに、近距離を撮影した観測画像を用いて処理を行った場合を考える。まず12枚の観測画像を用いた結果から、近距離撮影の欠点である視差や色合いの問題を考慮することで、問題なく超解像処理を行え、解像度を向上することができた (Figs.4.3.??)。

また、観測画像が少ない場合を比較してみると、従来手法 (Fig.4.7) ではエッジが強調されていると同時に、雑音も強調されてしまっていることがわかる。それに対し、提案手法 (Fig.4.8) では従来手法 (Fig.4.7) と同様にエッジが強調されているが、雑音は少ない状態であることがわかる。このことから、本研究で提案する手法は従来の超解像手法と同等以上の高解像度化が望め、雑音には強い手法であることがわかる。

また、S/N 比の比較 (Fig.4.29) から、提案手法は従来手法に比べて 2dB、超解像に用いた観測画像の枚数にして2枚分程良い結果が得られている。

次に、遠距離を撮影した画像を用いて処理を行った場合を考える。バイリニア法を用いて拡大した結果と比べてみると、多眼カメラを用いて撮影した画像に対して超解像処理が行われていることがわかる (Figs.4.9.4.10)。しかし、従来の手法を用いた結果と、提案する手法を用いた結果の視覚的に優位な差は確認できなかった (Figs.4.10.4.11)。観測画像が少ない場合を比較してみても、大きな視覚的な差は確認できなかった (Figs.4.12.4.13)。これは、観測画像内の雑音が少なかったことが大きな要因であると考えられる。このことは、それぞれの結果を拡大した画像 (Figs.4.14～.4.18) からよくわかる。また、S/N 比の比較 (Fig.4.30) から、従来手法と提案手法にあまり差が無いことがわかる。このS/N 比の差は観測画像の枚数が多くなるほど小さくなっている。

最後に、暗所を撮影した画像を用いて処理を行った場合を考える。バイリニア法を用いて拡大した結果と比べてみると、従来超解像処理結果の方が文字が読み取りやすくなっていることがわかる (Figs.4.19.4.20)。しかし、超解像処理結果は同時に雑音も強調してしまっている。また、従来手法 (Fig.4.20) と提案手法 (Fig.4.21) を比較してみると、どちらの手法も同じぐらい文字が強調されているが、従来手法に比べて提案手法の方が雑音が少ないことがわかる。観測画像が少ない場合も比較してみると、従来手法の画像には多くの雑音が残っているのに対し、提案手法の画像には雑音が少ないことがわかる (Figs.4.22.4.23)。これらの結果から、雑音低減を目的とした2つのフィルターを新たに加えた提案手法が上手く機能していると考えられる。また、S/N 比の比較 (Fig.4.31) から、提案手法は従来手法に比べて 2dB～3dB、超解像に用いた観測画像の枚数にして約3枚分程良い結果が得られている。

以上のことから、近距離撮影画像、暗所撮影画像を用いた提案手法による超解像処理は大幅に画質が改善でき、遠距離撮影画像を用いた場合は手法ごとにさほど差が表れないという結果になった。前者と後者の画像の主な違いとして、前者は後者に比べて観測画像内に多くの雑音が付加されているという点が挙げられる。前者の従来超解像処理結果、特に処理に用いた観測画像の枚数が少ない場合は画像内に雑音が残っている。それに対し、後者の従来超解像処理結果は画像内に残っている雑音が少ない状態にある。提案手法はこの雑音を取り除く、あるいは強調しないことができ、前者の提案手法超解像処理結果は従来手法に比べて大幅にS/N 比を向上できたのではないかと考えられる。また、すべての結果において雑音が強調されやすい観測画像の枚数が少ない場合の方が手法ごとのS/N 比の差が大きくなっ

ている。そのため、本研究で提案する手法は取得した観測画像内に雑音が多い、超解像処理に使用する観測画像の枚数が少ないなどの、超解像処理において雑音が表れてしまうような状況に強く、撮影枚数が少ない場合でも良い超解像処理の結果が得られるのではないかと考えられる。

第6章 まとめ

今回、本研究では多眼カメラのプロトタイプを作成し、様々な状態を撮影をして超解像処理を行った。提案する多眼カメラを用いた超解像処理として、(1) 観測画像に対しての色補正処理、(2) 視差が大きい場合の ROI 処理、(3) 観測画像の枚数が少ない場合のフィルター処理の3つの改善を取り入れて処理を行った。(1) と (2) により、各カメラごとのピントのずれなどによって色合いが違う場合や、撮影距離が近くて視差が大きくなってしまった観測画像に対しても問題なく超解像処理が行えた。また、(3) により、従来手法と比べて雑音に対して強い超解像処理が行え、平均して 2dB 程（観測画像約 2 枚分程度）の良い結果を得ることができた。今後は多眼カメラの小型化と共に、さらなる高解像度化を目指していきたい。

謝辞

本研究にあたり、全般にわたるご指導並びにデータセットの提供をしてくださった彌富仁准教授、および彌富研究室の皆様に深く御礼申し上げます。

参考文献

- [1] Park, S.C., Park, M.K., Kang, M.G.: Super-resolution image reconstruction: A technical overview. *IEEE Signal Process. Mag* 20 (2003) 21-36
- [2] Freeman, W.T., Jones, T.R., Pasztor, E.C.: Example-based super-resolution. *IEEE Computer Graphics and Applications* 22 (2002) 56-65
- [3] Takashi, K., Takahiro, S.: Super-resolution decoding of the jpeg coded image data using total-variation regularization. *Picture Coding Symposium* (2010) 114-117
- [4] Schultz, R., Stevenson, R.: A bayesian approach to image expansion for improved definition. *IEEE Transactions on Image Processing* 3 (1994) 233-242
- [5] Schultz, R., Stevenson, R.: Extraction of high-resolution frames from video sequences. *IEEE Transactions on Image Processing* 5 (1996) 996-1011
- [6] Zitova, B., Flusser, J.: Image registration methods: A survey. *Image and Vision Computing* 21 (2003) 977-1000
- [7] Tsai, R., Huang, T.: Multiframe image restoration and registration. *Advances in Computer Vision and Image* 1 (1984) 317-339
- [8] Eren, P., Sezan, M., Tekalp, A.: Robust, object-based high resolution image reconstruction from low-resolution video. *IEEE Transactions on Image Processing* 6 (1997) 1446-1451
- [9] Patti, A.J., Altunbasak, Y.: Artifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher-order interpolants. *IEEE Transactions on Image Processing* 10 (2001) 179-186
- [10] Zitova, B., Flusser, J.: Image registration methods: A survey. *Image and Vision Computing* 21 (2003) 977-1000
- [11] Bay, H., Ess, A., Tuytelaars, T., Gool, L.V.: Speeded-up robust features (surf). *Computer Vision and Image Understanding* 110 (2008) 346-359
- [12] 菅原正幸, 三谷公二, 齋藤敏紀, 藤田欣裕, 末次圭介, "4 板撮像方式における画素ずらし効果についての検討," *テレビジョン誌*, 49, 2, (1995) 212-218
- [13] Lowe, D.: Distinctive image features from scale invariant keypoints. *International Journal of Computer Vision* (2004) 91-110
- [14] Aghav, S., Kumar, A., Gadakar, G., Mehta, A., Mhaisane, A.: Mitigation of rotational constraints in image based plagiarism detection using perceptual hash. *International Journal of Computer Science Trends and Technology* 2 (2014) 28-32

- [15] Farsiu, S., Robinson, M.D., Elad, M., Milanfar, P.: Fast and robust multi-frame super-resolution. IEEE Transaction on image processing 13 (2004) 1327-1344

研究業績

国際学会

- Prototype of Super-Resolution Camera Array System, Daiki Hirao and Hitoshi Iyatomi, Lectures Notes in Computer Science, Vol. 9474, Vol. 1, pp.911-920, Dec. 2015. (International Symposium on Visual Computing 2015(ISVC2015)).

国内学会

- 多眼カメラを用いた超解像技術に関する基礎検討, 平尾 大樹, 彌富 仁, 2015 年電子情報通信学会総合大会 (滋賀) , 2015 年 3 月